

## ORIGINAL RESEARCH

**Unique secreted–surface protein complex of *Lactobacillus rhamnosus*, identified by phage display**Dragana Gagic<sup>1,\*</sup>, Wesley Wen<sup>1,a</sup>, Michael A. Collett<sup>2</sup> & Jasna Rakonjac<sup>1</sup><sup>1</sup>Institute of Molecular BioSciences, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand<sup>2</sup>Fonterra Research and Development Centre, Private Bag 11029, Palmerston North 4442, New Zealand**Keywords**Adhesins, bacteriophages, cell surface, *Lactobacillus***Correspondence**Jasna Rakonjac, Institute of Molecular BioSciences, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand.  
Tel: +6463505134; Fax: +6463505688;  
E-mail: j.rakonjac@massey.ac.nz**Funding Information**

This work was funded by the Institute of Molecular BioSciences (Massey University), Palmerston North Medical Research Foundation, Anonymous Donor, New Zealand Foundation for Research and Technology (contract number C03X0701) and Fonterra Co-operative Group Ltd. D. G. was a recipient of a Bright Future Enterprise Ph.D. Scholarship (Ministry of Science and Innovation, New Zealand).

Received: 17 October 2012; Accepted: 25 October 2012

***MicrobiologyOpen* 2013; 2(1): 1–17**

doi: 10.1002/mbo3.53

\*Current address: AgResearch Limited, Grasslands Research Centre, Private Bag 11008, Palmerston North 4442, New Zealand

<sup>a</sup>These two authors contributed equally to this study.**Introduction**

Bacterial surface proteins have been studied extensively in pathogenic bacteria over the last several decades, demonstrating their role in host colonization and cell invasion as well as very sophisticated manipulation of the host immune response pathways that often neutralize

**Abstract**

Proteins are the most diverse structures on bacterial surfaces; hence, they are candidates for species- and strain-specific interactions of bacteria with the host, environment, and other microorganisms. Genomics has decoded thousands of bacterial surface and secreted proteins, yet the function of most cannot be predicted because of the enormous variability and a lack of experimental data that would allow deduction of function through homology. Here, we used phage display to identify a pair of interacting extracellular proteins in the probiotic bacterium *Lactobacillus rhamnosus* HN001. A secreted protein, SpcA, containing two bacterial immunoglobulin-like domains type 3 (Big-3) and a domain distantly related to plant pathogen response domain 1 (PR-1-like) was identified by screening of an *L. rhamnosus* HN001 library using HN001 cells as bait. The SpcA-“docking” protein, SpcB, was in turn detected by another phage display library screening, using purified SpcA as bait. SpcB is a 3275-residue cell-surface protein that contains general features of large glycosylated Serine-rich adhesins/fibrils from gram-positive bacteria, including the hallmark signal sequence motif KxYKxGKxW. Both proteins are encoded by genes within a *L. rhamnosus*-unique gene cluster that distinguishes this species from other lactobacilli. To our knowledge, this is the first example of a secreted-docking protein pair identified in lactobacilli.

microbial-pattern-induced signaling pathways (Patti et al. 1994; Beckmann et al. 2002; Selbach and Backert 2005; Lilic et al. 2006; Luck et al. 2006; Timmer et al. 2006; Eskan et al. 2007; Gillen et al. 2008; Kulkarni et al. 2009; McGhie et al. 2009). Together with the diversity of nonprotein surface molecules and structures (cell wall peptidoglycan, lipoteichoic acid, and exopolysaccharides),

surface and secreted proteins determine the overall interaction of a pathogen with the host and the outcomes of the pathogen colonization (Grandel and Grimminger 2003; Boltana et al. 2011; Elberse et al. 2011; Hao et al. 2011; Tuanyok et al. 2012). Complex interactions with the host, including colonization and manipulation of immune responses, are not limited to pathogens. Commensal bacteria that colonize the gastrointestinal tract were found to be engaged in complex bacteria–host and bacteria–bacteria interactions that are mediated in part by surface and secreted proteins (Kolenbrander et al. 2006; Sansonetti and Medzhitov 2009; Sansonetti 2011). The advent of metagenomics brought appreciation for the diversity of microbial communities inhabiting the gastrointestinal system and other surfaces that are exposed to the environment (Grice and Segre 2011; Jenkinson 2011; Walter and Ley 2011; Dave et al. 2012), and understanding that the human gut–microbe interactions are in a state of immunological homeostasis in healthy individuals (Rimoldi et al. 2005; Sansonetti and Medzhitov 2009; Kinross et al. 2011). This homeostasis is disrupted when pathogens invade or expand, causing damage to the gut lining and profound inflammation (Sansonetti 2004). To restore the balance, preparations of live beneficial bacteria, called probiotics, are often recommended. Information on their role in the recovery from gut damage by pathogens and inflammation comes from animal models and clinical studies (Konstantinov et al. 2008; Preidis et al. 2011; Swanson et al. 2011; Wells 2011; Bron et al. 2012; Hummel et al. 2012). However, the surface or secreted proteins of the probiotic bacteria that likely mediate these beneficial effects are scarcely known (van Pijkeren et al. 2006; Kleerebezem et al. 2010; Wells 2011; Bron et al. 2012).

*Lactobacillus rhamnosus* is a species from the genus *Lactobacillus*, represented by several strains with demonstrated probiotic effects. The most prominent probiotic strains of *L. rhamnosus* are GG and HN001 (Gill et al. 2001; Sheih et al. 2001; Cross et al. 2002; Shu and Gill 2002; Lin et al. 2009; Thomas et al. 2011; Lahtinen et al. 2012). These two strains demonstrate aggregation in culture (Collado et al. 2007; Dekker et al. 2007). Both strains have been reported to have similar beneficial effects, including protection from enteric pathogens and prevention of eczema (Gill et al. 2001; Kalliomaki et al. 2003; Collado et al. 2006; Wickens et al. 2012).

A number of other lactobacilli species have been characterized as probiotics; some of those are isolates from dairy fermentation, whereas others are commensals (i.e., members of the resident gut microbial communities) (Wells 2011). Genome sequencing of a number of these species and strains have revealed a large number of diverse surface and secreted proteins, many of which are thought to have similar colonization roles as observed in bacterial

pathogens (van Pijkeren et al. 2006; Jankovic et al. 2007; Kleerebezem et al. 2010), such as mucin- and fibronectin-binding proteins (Toba et al. 1995; Altermann et al. 2005; Buck et al. 2005; Castaldo et al. 2009; Munoz-Provencio et al. 2010). Genomic information for a large number of poorly characterized commensal lactobacilli species is being collected within the United States and European human microbial genome projects (Dave et al. 2012).

Despite a large amount of genomic data on probiotic and commensal organisms, the functional understanding of extracellular (surface and secreted) proteins is far behind the model pathogenic bacteria (van Pijkeren et al. 2006; Kowalchuk et al. 2007; Song et al. 2009; Kleerebezem et al. 2010; Zhou et al., 2010; Kinross et al. 2011; Walter and Ley 2011). The high diversity of extracellular proteins means that their function, even in pathogenic bacteria, is not possible to identify by bioinformatics, in particular because the sequence similarity between proteins that have the same function is often limited to a genus, species, and even strain (e.g., proteins implicated in binding to the mucosal surfaces, or manipulation of immune response). These proteins therefore appear to represent recent evolutionary adaptations (Baltrus et al. 2011; Carroll et al. 2011).

One approach in identifying proteins that carry out a particular function, which cannot be derived from bioinformatics analyses, is to identify them by this function. This usually involves screening of expression libraries for activity or interactions. In this realm, a combinatorial affinity-selection method, called phage display (Zwick et al., 1998; Forrer et al. 1999; Sidhu et al. 2003; Rakonjac et al. 2011), has been successfully used to find extracellular proteins in pathogenic bacteria that interact with complex targets, such as bacterial or host cells, extracellular matrix proteins, or host antibody repertoire (Mullen et al. 2006). Phage display has also been used recently to identify the fibronectin-binding proteins from commensal bacterium *Lactobacillus casei* BL23 (Munoz-Provencio and Monedero 2011).

Cell-binding proteins of bacteria are often enzymes or proteins that mediate aggregation or biofilm formation. We used phage display to identify a cell-binding protein that interacts with the surface of HN001 cells (which we named SpcA), and its cognate “docking” cell-surface protein (named SpcB). The secreted protein SpcA has two bacterial immunoglobulin-like domains type 3 (Big-3) that mediate binding to the docking protein SpcB and a C-terminal domain of unknown function that is distantly related to plant pathogen response protein 1 (PR-1). The docking protein SpcB is the largest protein in *L. rhamnosus* genome; it has a standard LPXTG cell wall anchor motif, whereas its signal sequence and overall organization are characteristic of gigantic glycosylated Ser-rich proteins. Both proteins are unique to *L. rhamnosus*.

## Experimental Procedures

### Bacterial strains, growth conditions, and helper phage

*Escherichia coli* strain TG1 (*supE thi-1 Δ[lac-proAB] Δ[mcrB-hsdSM]5 [rK<sup>-</sup> mK<sup>-</sup>] [F' traD36 proAB lacI<sup>q</sup>ZΔM15]*) was used to construct and propagate all phagemid vectors and phage display library, and for obtaining stocks of the helper phage VCSM13 (Stratagene, La Jolla, California). TG1 containing the phagemid vectors or recombinant phagemids was infected with helper phage VCSM13 to allow packaging of phagemid DNA into phagemid particles (PPs) and display of *L. rhamnosus* proteins on the surface of the PPs. *Escherichia coli* cells were incubated in Yeast Extract Tryptone broth (2×YT) and *E. coli* transformants in 2×YT containing 20 μg/mL chloramphenicol (Cm) at 37°C with aeration. Lactobacilli were propagated in the Man–Rogosa–Sharpe (MRS) broth (Oxoid, Basingstoke, UK) at 37°C in liquid medium or on the Petri plates. Plates were incubated anaerobically. Bacteriological Agar (Oxoid) was used at 1% (w/v) for solid bacterial cultures. When required, antibiotics were added to media at the following concentrations: 20 μg/mL Cm and 100 μg/mL ampicillin (Amp) for *E. coli*; 5 μg/mL erythromycin (Em) for lactococci and lactobacilli; 5 μg/mL tetracycline (Tet) and 500 μg/mL streptomycin for lactobacilli. The detailed list of bacterial strains and their genotypes is given in Table S1.

### Phage infection, growth, purification, and titration

In the phagemid system used in this work, virions containing helper phage genomes are referred to as phage, whereas virions containing phagemid genomes are referred to as PPs. In all experiments in which phage or PPs were produced, an exponential-phase culture (around 10<sup>8</sup> cells/mL) of the appropriate strain was infected with helper phage VCSM13 at a multiplicity of infection (m.o.i) of 50 (50 phage to 1 bacterium), for 30 min at 37°C. Infected cells were separated from unabsorbed helper phage by centrifugation at 5000g for 10 min at room temperature and resuspended in fresh medium. The cultures were then incubated for 4 h at 37°C with aeration unless otherwise stated. At the end of incubation, cells were chilled, titrated (as required), and then separated from the phage or PPs by centrifugation at 13,200g for 20 min at 4°C. Bacterial pellets were resuspended in the equal volume of the chilled medium supplemented with 7% (v/v) DMSO (dimethyl sulfoxide). Generally, 1-mL aliquots were taken, frozen on dry ice, stored at –80°C, and later used for the analysis of the proteins by SDS-

PAGE (sodium dodecyl sulfate – polyacrylamide gel electrophoresis) and Western blotting.

The supernatant containing phages/PPs was centrifuged again and then passed through a 0.45-μm filter to eliminate any remaining bacterial cells. PPs were further purified and concentrated by precipitation in polyethylene glycol (PEG)/NaCl buffer (5% [w/v] PEG, 500 mmol/L NaCl) and resuspended in TN buffer (10 mmol/L Tris-HCl pH 7.6; 50 mmol/L NaCl) unless otherwise stated.

The number of infectious phage was determined by titration on TG1 strain. Approximate number of phage was first estimated by placing 10 μL drops of phage dilutions onto soft agar containing 0.2 mL of overnight culture of appropriate *E. coli* strain. Plaques in the area of absorbed drops were counted to determine approximate numbers of phage. For accurate titration of phage, dilution was adjusted to obtain 100–300 plaques per plate (based on the approximate titer). At least 300 plaques from the total of three plates were counted to determine the titer accurately. Helper phage titer was expressed as plaque-forming units per mL (pfu/mL).

The infectious PPs carrying a Cm<sup>R</sup> marker were titrated by the same method, except that specially prepared double-layer plates were used, to allow in-agar infection of the indicator strain TG1 prior to exposure to Cm (Russell 1993). Briefly, 2×YT agar plates were prepared with 25 μg/mL chloramphenicol. These plates were overlaid with 9-mL chloramphenicol-free 2×YT agar shortly before use. PP stock was then mixed with *E. coli* strain TG1 in the soft agar layer as described earlier. PP titer was expressed as colony-forming units per mL (cfu/mL).

### Construction of the phagemid pYW01 carrying PelB signal sequence and cold origin of replication

The phagemid pYW01 was constructed in two steps: first, the *psp* (filamentous phage infection induced) promoter, ribosomal-binding site and *pelB* signal sequence (from *Erwinia carotovora* pectate lyase – PelB) were amplified by polymerase chain reaction (PCR) using pJARA144 (J. Rakonjac, unpubl. data) as a template and primers pDJ01F01, containing an EarI site, and pJARA144R02, containing an NsiI restriction site (Table S2). The PCR product was cleaved and ligated into the EarI–NsiI digested phagemid pDJ01 (Table S2). This step inserted the *pelB* signal sequence between the *psp* promoter and multiple cloning site cassette of pDJ01. The resulting phagemid was designated pDJ08. Next, the origin of replication in pDJ08 (*colEI*) was replaced with the *colD* origin of replication from the vector pGZ119EH (Table S2). This was achieved by cleavage of the *psp-pelB-c-myc-gIIC* cassette and origin of replication from pDJ08 using PstI–ScaI

followed by removal of 3' protruding end from PstI by T4 DNA polymerase (Roche, Basel, Switzerland) and ligation of the resulting fragment to pGZ119EH EcoRV-ScaI fragment carrying *colD* origin of replication. The resulting phagemid pYW01 was used in the construction of *L. rhamnosus* HN001 shot-gun phage display library.

### Construction of the *L. rhamnosus* HN001 shot-gun phage display library

The library was constructed from mechanically (nebulization) sheared *L. rhamnosus* HN001 DNA and cloned into the phagemid vector pYW01. The fragments obtained varied in size between 0.3 and 4 kbp, with majority between 0.8 and 2 kbp. To eliminate fragments below 0.5 kbp, a protein concentrator with 100 kDa cutoff was used (Vivascience, Hannover, Germany).

Approximately 50  $\mu\text{g}$  of the genomic fragments were ligated to 14  $\mu\text{g}$  the SmaI-digested vector pYW01 using T4 DNA ligase (Roche). After phenol and chloroform extraction, the ligated DNA was ethanol precipitated, washed with 70% (v/v) ethanol, and dissolved in 70  $\mu\text{L}$  H<sub>2</sub>O. *Escherichia coli* TG1 electrocompetent cells (Jacobsson et al. 2003) were transformed with 3  $\mu\text{L}$  of ligation mix (2.5 kV, 25  $\mu\text{F}$ , 400  $\Omega$ ) in 2-mm-gap cuvettes. A total of 23 electrotransformations were carried out. The transformed cells were pooled, transferred to 230 mL of 2 $\times$ YT, and incubated for 1 h at 37°C with rotatory agitation. After incubation, a 1 mL aliquot was taken to determine the number of transformants by plating on 2 $\times$ YT agar with 20  $\mu\text{g}/\text{mL}$  Cm. An aliquot (10 mL) of the remaining bacteria was transferred to 990 mL of 2 $\times$ YT containing 20  $\mu\text{g}/\text{mL}$  Cm and incubated at 37°C with aeration until exponential phase (OD<sub>600</sub> 0.2). The exponentially growing culture was infected with wild-type helper phage VCSM13 (m.o.i = 50) for 30 min. Cells were then harvested by centrifugation at 3200g for 10 min and the pellet was resuspended in 1 L of 2 $\times$ YT containing 20  $\mu\text{g}/\text{mL}$  Cm. Infected cells were incubated for 4 h at 37°C with aeration. The host cells were pelleted by centrifugation and PPs were collected in the supernatant. These PPs represent the master phage display library. The library PPs were precipitated and purified as described in the phage manipulations section. They were stored in 7% (v/v) DMSO at -80°C.

### Affinity screening of the shot-gun phage display library for *L. rhamnosus* HN001 cells binders

*Lactobacillus rhamnosus* HN001-binding PPs were selected from the *L. rhamnosus* HN001 shot-gun library by affinity-enrichment protocol similar to the one described for the *Rhizobium leguminosarum* cell-surface-associated agglutinin

(Ausmees et al. 2001) with slight modification. Briefly, strain HN001 to be used as bait was cultivated in MRS medium supplemented with 1 mmol/L CaCl<sub>2</sub> at 37°C. The Ca<sup>2+</sup> ions were added based on the evidence that these ions are essential in binding of some proteins to the cell surface via lectin domains (Armitage et al. 1992; Buts et al. 2003). Cells from 1.5 mL of the late exponential phase ( $1 \times 10^9$  cells) were collected by centrifugation at 1100g for 5 min and the supernatant was carefully removed. The cell pellet was resuspended in 0.6 mL TBS (30 mM Tris, 150 mM NaCl, pH 7.6) containing 2 mmol/L CaCl<sub>2</sub> and mixed with the *L. rhamnosus* HN001 phage display library ( $3 \times 10^9$  PPs). Binding was allowed to proceed for 3 h at room temperature on a roller. Unbound PPs were removed by washing the cells five times with 1.5 mL TBS. After every wash, the bacteria were gently spun down at 1100g for 3 min in a microfuge. After the last wash, the cell pellet was resuspended in 0.05 mL 2 $\times$ YT, transferred to a new tube and mixed with 0.5 mL of the exponentially growing *E. coli* TG1 culture to recover bound phage. This mixture was incubated for 30 min at 37°C to allow phage to infect the *E. coli* cells. Subsequently, concentrated infected *E. coli* TG1 cells (and dilutions thereof) were plated on selective double-layer chloramphenicol plates (described above) to, respectively, recover and titer the eluted PPs. Cm<sup>R</sup> colonies were counted in diluted samples to determine the titer, whereas undiluted PPs were collected by scraping the plates with 2 mL 2 $\times$ YT. An aliquot (0.5 mL) was used to inoculate 200 mL of 2 $\times$ YT containing Cm. When the culture reached exponential phase, it was infected with wild-type helper phage VCSM13 in order to produce PPs of the enriched library. After the helper phage infection, the culture was processed, as described in the section on phage manipulation, to extract, concentrate, and purify PPs. These enriched PPs were then used in the subsequent round of panning according to the same protocol. Affinity-enrichment procedure was repeated three times (three rounds of panning were performed). After each round, the total number of eluted PPs was determined and profiles of plasmid dsDNA isolated from amplified library pools were monitored to estimate enrichment frequency. Distinct plasmid bands, enriched after the third cycle of panning, were purified from preparative agarose gel and used to transform the *E. coli* TG1 cells. Ten recombinants containing inserts of different sizes were further analyzed by DNA sequencing.

### Purification of PPs displaying SpcA and cell-binding assays

*Escherichia coli* clones containing the recombinant phagemid pSpcA<sup>A</sup> (K2089), selected from the HN001 shot-gun library after panning on *L. rhamnosus* HN001 cells



(encoding residues 1–376), and a clone containing recombinant phagemid pSpCA<sup>S</sup> (K2090) identified as a secretome clone in the HN001 secretome library (encoding residues 1–199) (Jankovic et al. 2007) were used to generate PPs displaying corresponding fragments of protein SpcA. A volume of 100 mL of exponentially growing cultures of corresponding strains K2089 and K2090 was infected with helper phage VCSM13 at m.o.i of 50 phages per bacterium. The negative PP control was pYW01 PPs, generated by infection of strain K1978 (TG1//pYW01) as described above. Infected cells were incubated for 4 h at 37°C with aeration. The *E. coli* cells were pelleted by centrifugation and PPs were collected in the supernatant. The PPs were precipitated with PEG/NaCl, resuspended in 200 µL of TBS pH 7.4, and enumerated by titration. Binding of SpcA<sup>A</sup>-displaying particles to HN001 cells did not require Ca<sup>2+</sup>; hence, all binding assays with individual clones were done in the absence of CaCl<sub>2</sub>.

To determine binding of PPs displaying SpcA<sup>A</sup>, and SpcA<sup>S</sup> fragments and binding of negative control containing empty vector pYW01, to HN001 and other tested strains, a binding assay that was equivalent to the panning protocol for library screening, described above, was used. The same protocol was used in binding assays to mutanolysin–lysozyme and proteinase-K-treated cells, except that the cells were treated with the appropriate enzymes and washed prior to assay, as described below.

For proteinase K treatment, HN001 cells from 1.5 mL of early stationary-phase culture were collected by centrifugation and washed once in TBS pH 7.4. The resulting cell pellet was then resuspended in 0.7 mL of buffer P(+) (100 mmol/L Tris-HCl pH 7.5, 5 mmol/L CaCl<sub>2</sub> and 5 mg/mL proteinase K) and buffer P(–) (100 mmol/L Tris-HCl pH 7.5 and 5 mmol/L CaCl<sub>2</sub>), respectively. All samples were incubated 1 h at 37°C. Cells were then washed four times in the equal volume of cell-wash buffer (TBS pH 7.4, 1 mmol/L EDTA [ethylenediaminetetraacetic acid], protease inhibitor cocktail [Roche]), by pelleting and resuspension, followed by a 15-min incubation at 15°C for all washes. After the last wash, cells were resuspended in 0.6 mL of the same buffer and used immediately in the binding assay as described above for untreated cells. Buffer control for this experiment contained the cells treated in the same manner, except that proteinase K was omitted. To ensure that protease did not degrade pIII or displayed SpcA on the surface of PPs during binding assay (which would result in loss of infectivity and/or failure to bind HN001 cells due to degradation of pIII and SpcA, respectively), a mock protocol was set up using a mixture of cell-wash buffer (TBS pH 7.4, 1 mmol/L EDTA, protease inhibitor cocktail), and the proteinase-K-containing buffer, allowing for expected protease dilution in the PPs-cell-binding step. The titer

remained the same, confirming no degradation of pIII. For lysozyme–mutanolysin (L-M) treatment, HN001 cells from 20 mL of late-exponential-phase culture were collected by centrifugation and resuspended in equal volume of prewarmed MRS broth. After the culture was incubated for another 2 h, cells from 1.5 mL of culture were collected and washed once in TBS pH 7.4. The cells were then resuspended in 0.5 mL of L-M buffer(+) (TBS pH 7.4, 25% [w/v] sucrose, 1 mmol/L EDTA, 20 mg/mL lysozyme, and 20 µg/mL mutanolysin) and L-M buffer(–) (TBS pH 7.4, 25% [w/v] sucrose and 1 mmol/L EDTA), respectively. After reaction was carried out at 37°C for 1 h, 0.5 mL of 0.25 mol/L EDTA was added. The mixture was incubated on ice for 5 min to stop the reaction. The resulting protoplasts were collected by gentle centrifugation (1000g), washed once with TBS, and resuspended in 0.6 mL of the assay buffer (TBS). Buffer control for this experiment contained the cells treated in the same manner, except that lysozyme and mutanolysin were omitted.

Binding efficiency in all cell-binding assays was determined as the ratio of total number of eluted PPs, to the total number of PPs used in the assay. The number of particles was determined by titration, as described in the phage manipulations section.

### Expression and purification of MBP-SpcA fusion protein

DNA sequence encoding the N-terminal portion of SpcA, containing both Big-3 domains, but lacking the signal sequence (SpcA<sup>A1</sup>, residues 117–309) was amplified by PCR, using genomic DNA as template and primers pWW37 and pWW50 (Table S3). The primers introduced EcoRI and PstI sites at the ends of the amplicon. The PCR product and the expression vector pMal-c2X (New England Biolabs, Ipswich, Massachusetts) were digested with EcoRI and PstI restriction enzymes and ligated to EcoRI-PstI-cut pMAL-c2X vector which contains maltose-binding protein (MBP) N-terminal tag to obtain recombinant plasmid pMBP-SpcA<sup>A1</sup>. The MBP-SpcA<sup>A1</sup> fusion was then expressed in TG-1 cells and affinity purified using amylose resin (New England Biolabs) according to the manufacturer's instructions. Purity of the protein was analyzed by SDS-PAGE (Laemmli 1970).

### Screening of *L. rhamnosus* HN001 shot-gun phage display genomic library for gene(s) encoding SpcA-binding protein(s)

Purified MBP-SpcA<sup>A1</sup> fusion was used as a bait to screen the HN001 shot-gun phage display library (described above) to identify SpcA-docking protein on the surface of HN001. Four rounds of panning were carried out. In the first round

of panning, affinity binding to amylose resin (binds MBP tag of the MBP-SpcA<sup>A1</sup> fusion with a high affinity; New England Biolabs) was used in the enrichment step of the library panning. Briefly, 4  $\mu$ g MBP-SpcA<sup>A1</sup> fusion was mixed with 10<sup>11</sup> PPs from the library in a microcentrifuge tube. After incubation at room temperature for 3 h, 50  $\mu$ L of amylose resin in TBS was added and the mixture was incubated for another hour. The amylose resin with bound proteins and phage particles was then collected by centrifugation and washed five times with TBS to eliminate unbound proteins and PPs. After the last wash, the bound PPs and MBP-SpcA<sup>A1</sup> protein were eluted by resuspending the amylose resin in 500  $\mu$ L of TBS containing 10 mmol/L maltose at room temperature for 15 min. Phage particles in complex with MBP-SpcA<sup>A1</sup> fusion protein were amplified by infecting *E. coli* TG-1 to transduce the phagemid DNA; the transfected (Cm<sup>R</sup>) cells were then superinfected with VCSM13 helper phage to produce the PPs, which are concentrated and purified as described above for the library screening using HN001 cells as a bait. In the next three rounds of panning, the affinity-enrichment procedure was carried out in the microtiter plates (MaxiSorp, Nunc, Rochester, New York). The bait protein (MBP-SpcA<sup>A1</sup>) and controls (MBP and bovine serum albumin [BSA]) were immobilized on the surface of microtiter plate wells at 4°C overnight, and the wells were subsequently blocked by 2% (w/v) BSA in TBS. About 3  $\times$  10<sup>9</sup> phage particles from the previous panning round on MBP-SpcA<sup>A1</sup> bait were added. After incubation at and washing, MBP-SpcA<sup>A1</sup>-bound phage particles were eluted by 100  $\mu$ L of low-pH elution buffer (100 mmol/L glycine-HCl, 0.1% [w/v] BSA, pH 2.2) at room temperature for 15 min and neutralized by 6  $\mu$ L of 2 mol/L unbuffered Tris. Eluted PPs were amplified and concentrated, and the plasmid profiles after each round of panning were examined as described in the section of library screening for binding to HN001 cells. The profile of plasmids eluted from parallel controls, panned on BSA and MBP, was used as a comparison, to distinguish plasmid bands selected on the MBP-SpcA<sup>A1</sup> fusion relative to those selected by panning on the controls. One plasmid band that was identified in the plasmid profile after the fourth round of panning on MBP-SpcA<sup>A1</sup>, but absent from the profiles of plasmids eluted from the same round of panning on negative controls (BSA and MBP), was excised from the gel and transformed into TG1. Inserts of ten individual transformants were analyzed by DNA sequencing to identify open reading frames (ORFs) from which they are derived.

### Phage-based Western blot

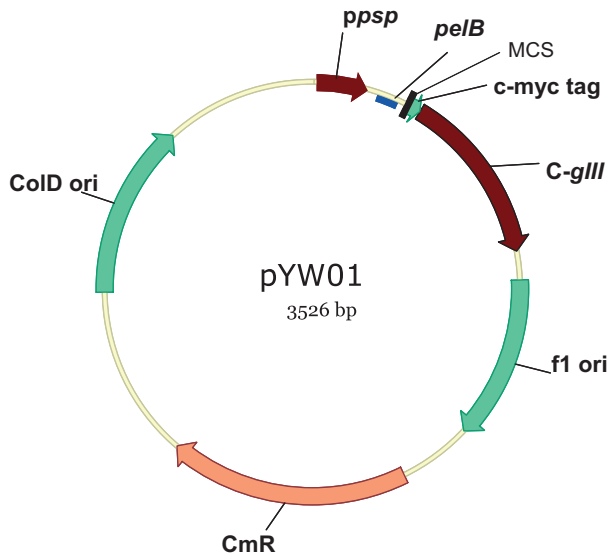
To confirm binding of PPs displaying SpcB<sup>N1</sup> and SpcB<sup>N2</sup> fragments of protein SpcB to the SpcA<sup>A1</sup>, phage-Western blots (or PP-Western blots) were carried out. The MBP-

SpcA<sup>A1</sup> fusion protein was partially cleaved by the sequence-specific protease, Factor Xa (Novagen-Merck, Darmstadt, Germany), whose recognition site is between the MBP tag and the SpcA protein in pMBP-SpcA<sup>A1</sup>. The Factor Xa digestion reactions were prepared according to the product manual (New England Biolabs) and incubated at 37°C for 2 h. Two sets of digested proteins and undigested controls were analyzed by SDS-PAGE; gel portion containing one set of samples was stained by Coomassie brilliant blue and the portion containing a duplicate set was transferred to the nitrocellulose membrane for blotting with SpcB<sup>N1</sup>- and SpcB<sup>N2</sup>-displaying PPs. The membranes were blocked in TBST (TBS; 0.5% [v/v] Tween-20) containing 5% (w/v) skim milk at 4°C overnight. The membranes were subsequently exposed to 1  $\times$  10<sup>9</sup> phage particles in TBST at room temperature for 2 h and the unbound phage particles were washed off with TBS. The bound phage particles were in turn detected by primary antibodies against M13 phage coat protein pVIII (Sigma-Aldrich, St. Louis, Missouri) and alkaline phosphatase-conjugated secondary antibodies using the standard Western blot protocols. The PP-SpcB<sup>N1</sup> and PP-SpcB<sup>N2</sup> – bound bands on the membrane – were detected by alkaline phosphatase assay (Sambrook and Russell 2001).

## Results

### Construction of a new phage display vector, pYW01, and a shot-gun phage display library of *L. rhamnosus* HN001

A phage display phagemid vector suitable for bacterial shot-gun phage display libraries was constructed, to avoid toxicity of bacterial proteins to *E. coli* and improve display efficiency. The display cassette is expressed from a phage-shock-protein (*psp*) promoter, induced by the helper filamentous phage infection. The *psp* promoter is only expressed after the helper phage infection and was shown to minimize toxicity of pIII fusions that is often encountered in phage display (Beekwilder et al. 1999). The display cassette of this vector (named pYW01) contains the PelB signal sequence, a multiple cloning site, followed by a c-myc tag fused to the C-terminal domain of pIII (pIIIC; Fig. 1). The truncated pIII (pIIIC) in pYW01 lacks the N-terminal domains of pIII which cause resistance to filamentous phage infection when expressed in host *E. coli* cells (Rakonjac et al. 2011). This is important in construction of bacterial genomic shot-gun phage display libraries, to avoid resistance to infection with the helper phage (required for production of PPs and display of proteins encoded by the inserts) when the inserts contain bacterial promoters upstream of the ORFs encoding displayed proteins.



**Figure 1.** Phagemid vector used in construction of *Lactobacillus rhamnosus* HN001 shot-gun phage display library. Vector features (in the clockwise direction) are *ppsp*, phage shock protein promoter; *pelB*, encoding a PelB protein signal sequence; MCS, multiple cloning cassette; c-myc tag, a common peptide tag; *C-gIII*, C-domain of *gIII*; f1 ori, the f1 phage origin of replication for generation of ssDNA for packaging into the phagemid particles; CmR, chloramphenicol resistance cassette; colD ori, the colD plasmid origin of replication.

A shot-gun phage display library of HN001 was constructed in this new vector, as described in Experimental Procedures. The inserts of the library ranged from 0.3 to 4 Kbp; the primary size of the library was  $3 \times 10^8$  clones. The library was first amplified using the plasmid origin of replication (in the absence of a helper phage). In the next step, the amplified library was mass infected with the *gIII*<sup>+</sup> helper phage VCSM13 to initiate replication from the phagemid f1 origin and packaging into PPs. These PPs were expected to display the whole HN001 proteome and contain corresponding DNA sequences inside the virions. PPs released from the infected library were collected, purified by PEG precipitation, and used in affinity screenings.

### **SpcA is a cell-binding protein of *L. rhamnosus* HN001**

Cell-surface-binding proteins in bacteria may be involved in a number of functions, including cell aggregation, or they can be surface-associated enzymes or adhesins. We undertook to functionally select for cell-surface-binding proteins by affinity screening of an HN001 phage display library for recombinant PPs that bind the HN001 cells.

For the enrichment procedure, strain HN001 cells were used as bait to enrich for cell-adhering phagemid clones

from the library. The library enrichment for PPs that bind to HN001 cells was monitored by increase in the number of eluted PPs after each progressive round of panning. A 916-fold increase in the number of eluted PPs between rounds one and three indicated that specific enrichment had taken place. Plasmid profiles of enriched library pools after each round of panning were monitored by agarose gel electrophoresis as another indicator of enrichment for selected clones (Data S1 and Fig. S1). Before the panning, the random size distribution of inserts resulted in a smear upon agarose electrophoresis of purified plasmids, whereas after the second and third round discrete plasmid bands were observed, indicating enrichment for plasmids of specific sizes. The most prominent three bands seen on the agarose gel after the third panning round were each excised, purified, and used to separately transform *E. coli* host strain TG1. Sequence analyses of inserts from ten transformants revealed that all contained the 5'-moiety of a single ORF of a secretome protein, in frame with the c-myc tag and *gIII*C. The ORF sequence was preceded by a putative ribosomal-binding site and initiated by an ATG start codon. Upstream of the transcriptional site a -10 box (ACATAAAAT) and -35 box (TTGATT) was identified in the putative promoter region using BPROM promoter prediction program from the Softberry server (<http://www.softberry.com/berry.phtml>). Translated products of the inserts encoding the selected ORF ranged in size from 334 to 376 amino acids. Searching of an *L. rhamnosus* HN001 genome sequence identified this ORF as *lrh\_07061* and its product is annotated as a cell-surface protein precursor (accession number ZP\_03210506 within the contig NZ\_ABWJ01000002.1). As the first *L. rhamnosus* protein whose cell-binding properties are experimentally proven, the locus and the gene were renamed *spcA* (surface protein, cell-binder A).

The signal sequence prediction algorithm SignalP 4.0 (Petersen et al. 2011) revealed that the *spcA* encodes an extracellular protein with an amino-terminal signal sequence. Furthermore, Pfam, CDART, CDD, and SMART database searches (Schultz et al. 1998; Geer et al. 2002; Finn et al. 2006; Marchler-Bauer et al. 2011) detected possible functional domains, two bacterial immunoglobulin-like type 3 domains (Big-3; E-values  $4e^{-04}$  and  $3e^{-05}$ ). The Ig-like domains are commonly found in the surface and secreted proteins in both prokaryotes and eukaryotes and are typically involved in extracellular adhesion and binding processes (Hultgren et al. 1993; Takai and Ono 2001; Fraser et al. 2007).

The translated complete ORF of *spcA* was 503 residues in length (1512 bp). In addition to the Big-3 domains, SpcA in all sequenced *L. rhamnosus* strains contains a third domain that is C-terminal to the Big-3 domains. This domain was identified as an SCP-like extracellular

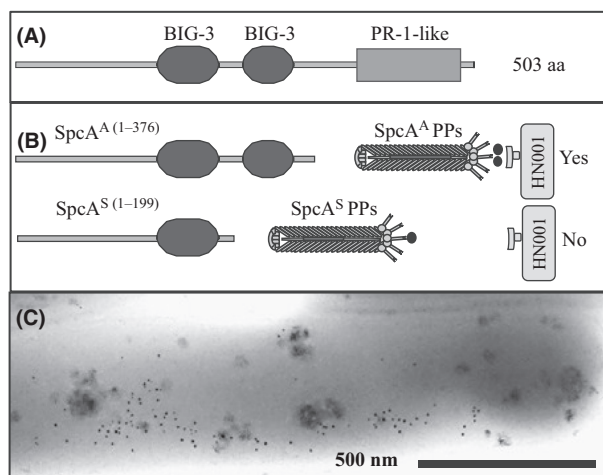
protein domain (SCP\_bacterial; CD05379;  $E = 4e^{-3}$ ). The SCP protein was initially identified as a murine extracellular glycoprotein (Mizuki and Kasahara 1992), and the SCP-like domains are found in diverse extracellular proteins of bacteria and eukaryotes. Some members of this family are endopeptidases, while some are transglycosylases or have nonenzymatic roles. Analysis using the PHYRE protein structure prediction web server (Kelley and Sternberg 2009) selected a member of this family from plants, pathogen-recognition protein PR-1-like, as the top structural homolog along the whole length of this domain; therefore, this domain in SpcA was denoted as PR-1-like (Fig. 2A). The equivalent ORF to SpcA in the genome sequence of *L. rhamnosus* GG is denoted as LGG\_01989 (Kankainen et al. 2009). Interestingly, although the SpcA amino acid sequence over 503 residues is 100% and 97% conserved with those of *L. rhamnosus* strains GG and Lc 705, respectively, in the two latter strains this ORF has additional 68 residues at the C-terminus, whereas the HN001 ORF is disrupted by an IS5-like element.

A secretome database/clone bank, a resource of displayed purified surface and secreted proteins from *L. rhamnosus* HN001 constructed previously in our laboratory (Jankovic et al. 2007), contained a clone that displayed a truncated spcA (ACC61714) that encoded a

199-residue fragment. This fragment was shorter than inserts selected for binding to HN001 cells; it encoded only one (N-terminal) Big-3 domain. To distinguish between the shorter SpcA insert from the clone bank, and a longer affinity-selected larger insert, the shorter one was named SpcA<sup>S</sup> (for short), whereas the longer one was named SpcA<sup>A</sup> (for affinity selected). The short fragment, which encoded only the N-terminal Big-3 domain, was used for mapping of the HN001-binding portion of SpcA. PPs displaying this N-terminal 199-residue long SpcA<sup>S</sup> fragment showed no binding above the vector background (Fig. 2). Therefore, for interaction with the HN001 surface structures, sequence downstream of the N-terminal Big-3 domain, containing the C-terminal Ig-like domain, is required. Binding of the SpcA<sup>A</sup> to the cells was confirmed by transmission electron microscopy which detected gold-labeled SpcA<sup>A</sup>-displaying PPs bound to the surface of HN001 cells (Fig. 2C).

When a BLASTP search (Altschul et al. 1997) was performed with the complete amino acid sequences from HN001 SpcA or its GG homolog, a hit with low conservation to the second (C-terminal) Big-3 domain of SpcA ( $E = 5e^{-3}$ ) but with a highly conserved PR-1-like domain and downstream 68-residue extension found in GG but not in HN001 ( $E = 6e^{-67}$  to HN001 and  $E = 1e^{-85}$  to GG) was identified (*Lactobacillus zeae* KCTC 3804; accession number ZP\_09454360.1). Other hits had a relatively low significance; the homologies were limited to the pair of Big-3 domains and were not secreted, but are rather LPXTG cell-wall-anchored proteins found in genera *Lactobacillus*, *Enterococcus*, and *Listeria*.

Together, these observations indicated that the *spcA* ORF encodes a *L. rhamnosus* – unique secreted protein with a propensity for multiple interactions through two Big-3 domains and one PR-1-like domain.

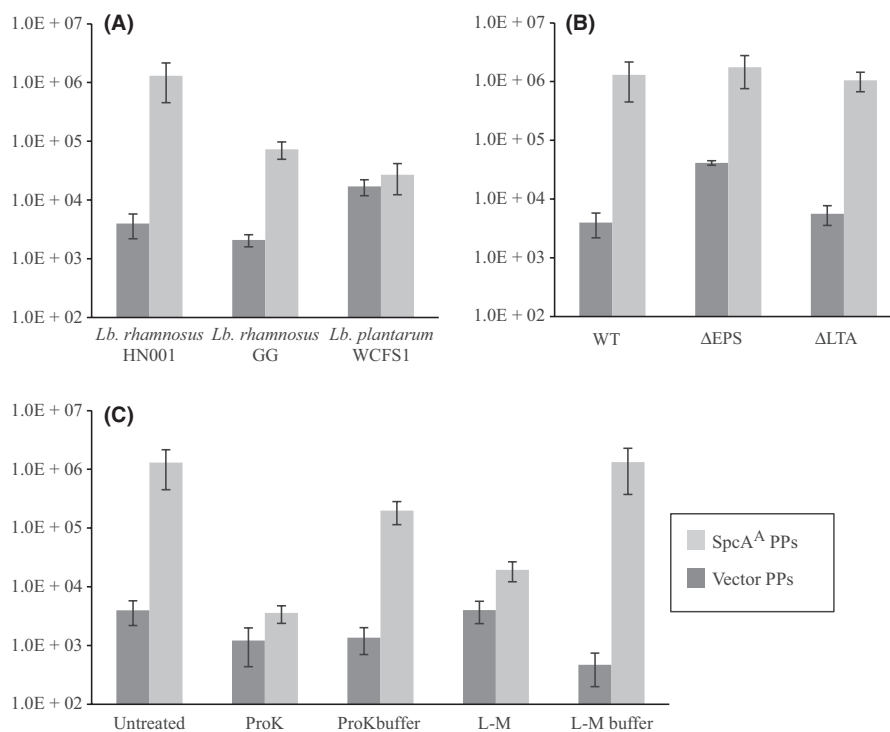


**Figure 2.** SpcA protein domain organization and mapping of HN001-binding domains. (A) SpcA protein domain organization. SpcA protein consists of two bacterial immunoglobulin-like domains family 3 (Big-3) and a putative PR-1-like domain. (B) Mapping of HN001-binding domains. The phagemid particles (PPs) displaying both Big-3 domains (SpcA<sup>A</sup> PPs) were capable of binding to HN001 cells, whereas the phage particles displaying only N-terminal Big-3 domain (SpcA<sup>S</sup> PPs) lost the binding capacity. The numbers in parentheses indicate the length of displayed fragment (in amino acid residues). (C) Electron micrograph of gold-labeled SpcA<sup>A</sup> PPs attached to the surface of an HN001 cell.

### SpcA binds to a cell-wall-associated protein unique to *L. rhamnosus* species

To determine whether SpcA has a role in cell–cell aggregation, a deletion of *spcA* gene was constructed by marker exchange, as described in Figure S2. The  $\Delta$ *spcA* mutant did not demonstrate any difference from the wild type in aggregation (not shown), suggesting that this is not an aggregation protein. Given that the SCP-like or PR-1-like domains in other organisms are versatile and have been reported to perform a variety of functions, from enzymatic to host-defense, we hypothesized that the SpcA Big-3 domains may bind to the bacterial cell in order to immobilize the PR-1-like domain on the surface of the cell. In that case, it is expected that SpcA would be anchored through a cognate species-specific docking molecule. In order to determine the SpcA tropism or





**Figure 3.** Lrh33 cell-binding assays. In these assays, binding of SpcA<sup>A</sup>-displaying phagemid particles (PPs) that were selected from HN001 phage display library (SpcA<sup>A</sup> PPs; Fig. 2) was compared with binding of the negative control PPs derived from empty vector pYW01 (Vector PPs) as described in Experimental Procedures. (A) Binding assays with different *Lactobacillus* strains and species. (B) Binding assays with the wild-type (WT) *L. rhamnosus* HN001 and mutants; ΔEPS, mutant lacking an extracellular exopolysaccharide; ΔLTA, mutant with an altered lipoteichoic acid. (C) Binding assays with lysozyme–mutanolysin (L-M) treated HN001 and proteinase K treated HN001 cells. Untreated, HN001 cells in phosphate buffer saline (PBS); ProK, proteinase-K-treated cells; ProK buffer, cells processed in the same manner as in the ProK assay, except that the proteinase K was omitted; L-M, lysozyme–mutanolysin-treated cells (spheroplasts); L-M buffer, cells processed as L-M cells, except that the lysozyme and mutanolysin are omitted.

species specificity, we tested the binding profile of SpcA<sup>A</sup>-displaying PPs (SpcA<sup>A</sup> PPs) to various lactobacilli. SpcA<sup>A</sup> showed binding to *L. rhamnosus* GG, but not to different lactobacillus species, *Lactobacillus plantarum* strains WCFS1 (Fig. 3A) and ATCC14197, *Lactobacillus sakei* Lb790 and *Lactobacillus acidophilus* NCFM (data not shown). Therefore, Lrh33 binds to a ligand that is most likely species specific for *L. rhamnosus*.

Next, we sought to determine the nature of the SpcA-docking molecule on the surface of *L. rhamnosus* cells. We tested binding of SpcA<sup>A</sup> PPs to the mutant HN001 lacking an extracellular polysaccharide or with an altered lipoteichoic acid (LTA), but these mutations did not affect binding of SpcA<sup>A</sup> PPs (Fig. 3B). To test whether SpcA<sup>A</sup> PPs interact with the surface proteins, their binding to the proteinase-K-treated cells was tested (Fig. 3C). Proteinase K eliminated binding of SpcA<sup>A</sup> PPs. Furthermore, SpcA<sup>A</sup> PPs did not bind to spheroplasts, which were obtained by L-M treatment of HN001. These enzymes removed the cell wall, together with the cell-

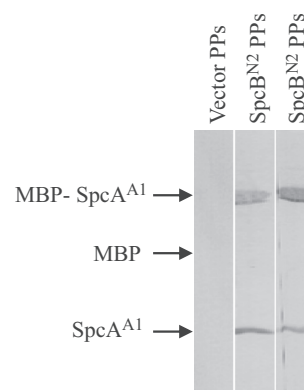
wall-anchored proteins, from the surface of the cells. In summary, lack of binding to protease-treated cells and spheroplasts and binding to cells lacking an exopolysaccharide (EPS) or with an altered LTA suggests binding to a cell-wall-anchored protein.

### SpcB, the largest surface protein in *L. rhamnosus* genome, is the SpcA-docking protein

To identify SpcA ligand(s), we undertook another affinity screening of the HN001 phage display library, and this time using as bait purified MBP fusion to the SpcA<sup>A</sup> fragment that lacks signal sequence (residues 117-309; named MBP-SpcA<sup>A1</sup>). The first round of panning was carried out by immobilizing the MBP-SpcA<sup>A1</sup>, in complex with associated library PPs, on the amylose resin which captures the MBP tag fused to SpcA<sup>A1</sup> fragment. The following three rounds of panning were performed using MBP-SpcA<sup>A1</sup> fusion immobilized on the microtiter plate.

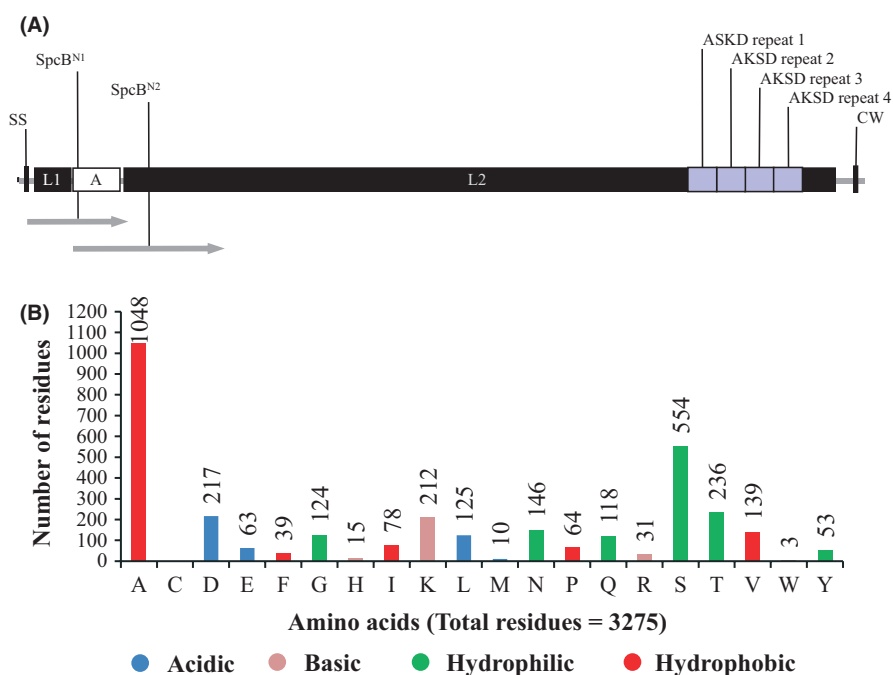
In the rounds 2–4, the library was first depleted from MBP-binding clones by first binding to immobilized MBP tag; the remaining unbound PPs were then used for binding the MBP-SpcA<sup>A1</sup> fusion. Washing and elution steps were performed for both the MBP and MBP-SpcA<sup>A1</sup> and plasmid profiles were analyzed for both eluates, to distinguish between SpcA<sup>A1</sup>- and potential MBP-binding PPs. Analysis of the plasmid profiles after the fourth round of panning indicated enrichment of discrete plasmid bands. The bands that were present in the eluate from the immobilized MBP-SpcA<sup>A1</sup> but absent from the MBP control eluate (Fig. S3) were excised from agarose gels and transformed into *E. coli* TG1 strain to obtain individual clones. Ten transformants were analyzed by DNA sequencing; the inserts in five clones encoded the N-terminal domain of SpcB, the largest protein in the HN001 genome (3275 residues in length). This protein is encoded by *spcB* gene, located in the same cluster of *L. rhamnosus*-unique genes with *spcA*. These five transformants represent two clones containing distinct but overlapping inserts, named SpcB<sup>N1</sup> and SpcB<sup>N2</sup>, encoding amino acid residues 32–391 and 211–586, respectively. Both inserts were in frame with upstream vector-encoded PelB signal sequence and downstream c-myc tag and C-terminal domain of pIII, confirming that they were able to be displayed on the surface of PPs. To confirm binding to SpcA<sup>A1</sup>, rather than the MBP portion of the bait, the MBP-SpcA<sup>A1</sup> fusion was digested with the sequence-specific protease Factor Xa, whose cleavage site is engineered between the MBP and SpcA<sup>A1</sup>. Binding of SpcB<sup>N1</sup>- and SpcB<sup>N2</sup>-displaying PPs to either of the proteolytic products was determined by a phage-Western blot. This is a blot in which the partially digested MBP-SpcA<sup>A1</sup> fusion, separated by SDS-PAGE and transferred onto a nitrocellulose membrane, is probed with the SpcB<sup>N1/2</sup>-displaying PPs (and the empty vector control PPs), followed by detection of bound phage using antiphage antibodies. This experiment showed specific binding of the SpcB<sup>N1</sup> PPs and SpcB<sup>N2</sup> PPs to SpcA<sup>A1</sup> and to the uncut MBP-SpcA<sup>A1</sup> fusion, but not to the MBP band, proving that this is specific SpcA<sup>A1</sup>-SpcB<sup>N</sup> interaction (Fig. 4). The sequence overlap between the two inserts, SpcB<sup>N1</sup> and SpcB<sup>N2</sup>, determines an SpcA<sup>A1</sup>-binding domain, from residue 211–391.

The SpcB primary sequence organization (Fig. 5) resembles that of large glycosylated Ser-rich surface adhesins of Gram-positive bacteria, including the characteristic signal sequence that contains KxYKxGKxW motif within an extended N-terminal positively charged sequence, and a glycine-rich hydrophobic region (Bensing et al. 2007). The high-complexity SpcA-ligand-binding domain (A region) is preceded by a relatively short Ala-Ser-Thr-rich low-complexity region (L1; residues 67–213) at the N-ter-



**Figure 4.** Confirmation of the physical interaction between SpcA and SpcB by phagemid particles (PPs) Western blot. Affinity-purified MBP-SpcA<sup>A1</sup> was first partially cleaved by a protease, Factor Xa, which cuts between the MBP and SpcA<sup>A1</sup>. Uncleaved fusion protein and proteolysis products, MBP and SpcA<sup>A1</sup>, were separated by SDS-PAGE and transferred to nitrocellulose membranes. The membranes were cut along the lanes in the gel, and each strip was blotted with one of the three PP samples: Vector pYW01 PPs, PPs containing empty vector (negative control); SpcB<sup>N1</sup> PPs, PPs displaying SpcB<sup>N1</sup> (residues 59–391); SpcB<sup>N2</sup> PPs, PPs displaying SpcB<sup>N2</sup> (residues 211–595); MBP, maltose-binding protein N-terminal tag. The bound PPs were detected using primary rabbit anti-M13 phage antibodies, followed by anti-rabbit IgG secondary antibodies conjugated with alkaline phosphatase, and visualized by alkaline phosphatase enzymatic assay.

minus of the mature protein and is followed by a nearly 3000-residue long low-complexity Ala-Ser-rich domain (L2; residues 389–3178) in *L. rhamnosus* HN001 and GG (Kankainen et al. 2009; Morita et al. 2009). As a difference from other Gram-positive Ser-rich proteins, SpcB does not have a short highly conserved Ser-rich motif; instead, it is organized as imperfect repeats of about 100 residues in length. Four almost identical 111-residue repeats are located between residues 2591 and 3034; these repeats are more hydrophobic than the upstream low-complexity region. Besides the predominant Ala (44) and Ser (20) residues, these repeats contain a large number of charged residues, mainly Lys (14) and Asp (13), with other residues having low representation and were therefore named ASKD repeats. Relatively high frequency of these four residues is maintained along the whole low-complexity region which makes up most of the protein (Fig. 5B). Along the whole length of the low-complexity region, the PHYRE server predicts both the disordered structure (with high probability) and  $\alpha$ -helical secondary structure (with variable probability). It is therefore possible that this part of the molecule forms a helical fibril. It is noteworthy that the SpcB sequence between residues 401 and 1209, when fused to a signal-sequence-less pIII, was able to target the fusion to the *E. coli* membrane (Jankovic et al. 2007; Jankovic 2009); therefore, this region of SpcB is likely to be



**Figure 5.** SpcB protein organization and mapping of SpcA-binding region. (A) Schematic representation of the SpcB organization. SS, the KxYKxGKxW signal sequence; L1, low-complexity region I; A, high-complexity region containing SpcA-binding sequence; L2, low-complexity region II; CW, cell-wall anchor, including LPXTG motif. The gray block arrows represent the two N-terminal SpcB fragments that were affinity selected from the HN001 phage display library using purified MBP-SpcA<sup>A1</sup> protein as bait. (B) Plot of the SpcB amino acid composition.

able to insert into lipid bilayers. The C-terminal LPXTG motif (Schneewind et al. 1992) and transmembrane helix of SpcB correspond to a typical cell wall anchor. Finding that the ligand is a *L. rhamnosus*-unique cell-wall-anchored protein is consistent with the SpcA<sup>A</sup> PP binding assays described in the previous section. In GG, as in HN001, SpcB is 3275 residues in length, whereas in Lc 705 it has 3390 residues, due to an extra 111-residue repeat within the Ala-Ser-rich region, and several other short insertions/duplications.

Whereas no homologs were identified for the Ala-Ser-rich sequence which represents the majority of the protein (when the BLASTP search was carried out with the low-complexity filter turned on), the high-complexity SpcA-binding domain had low similarity to one protein in the NCBI database, a hypothetical KxYKxGKxW-signal sequence protein of the human gut bacterium *Lactobacillus parafarraginis* F0439 (ZP\_09393738.1). This hypothetical protein has 32% identity (BLASTP;  $E = 3e^{-12}$ ) to SpcB between residues 233 and 449, corresponding roughly to the N-terminal high-complexity domain that mediates binding to SpcA Big3 domains; furthermore, the two proteins have a conserved signal sequence (residues 1–60; 57% identity;  $E = 2e^{-10}$  by BLAST analysis). The SpcA-binding domain is not annotated in the pfam database and did not give significant hits in the PHYRE

three-dimensional (3D) alignment, and hence, it may be considered a unique domain. The remaining portion of *L. parafarraginis* ZP\_09393738.1 protein also has a low-complexity Ala-Ser-rich region arranged in imperfect repeats.

## Discussion

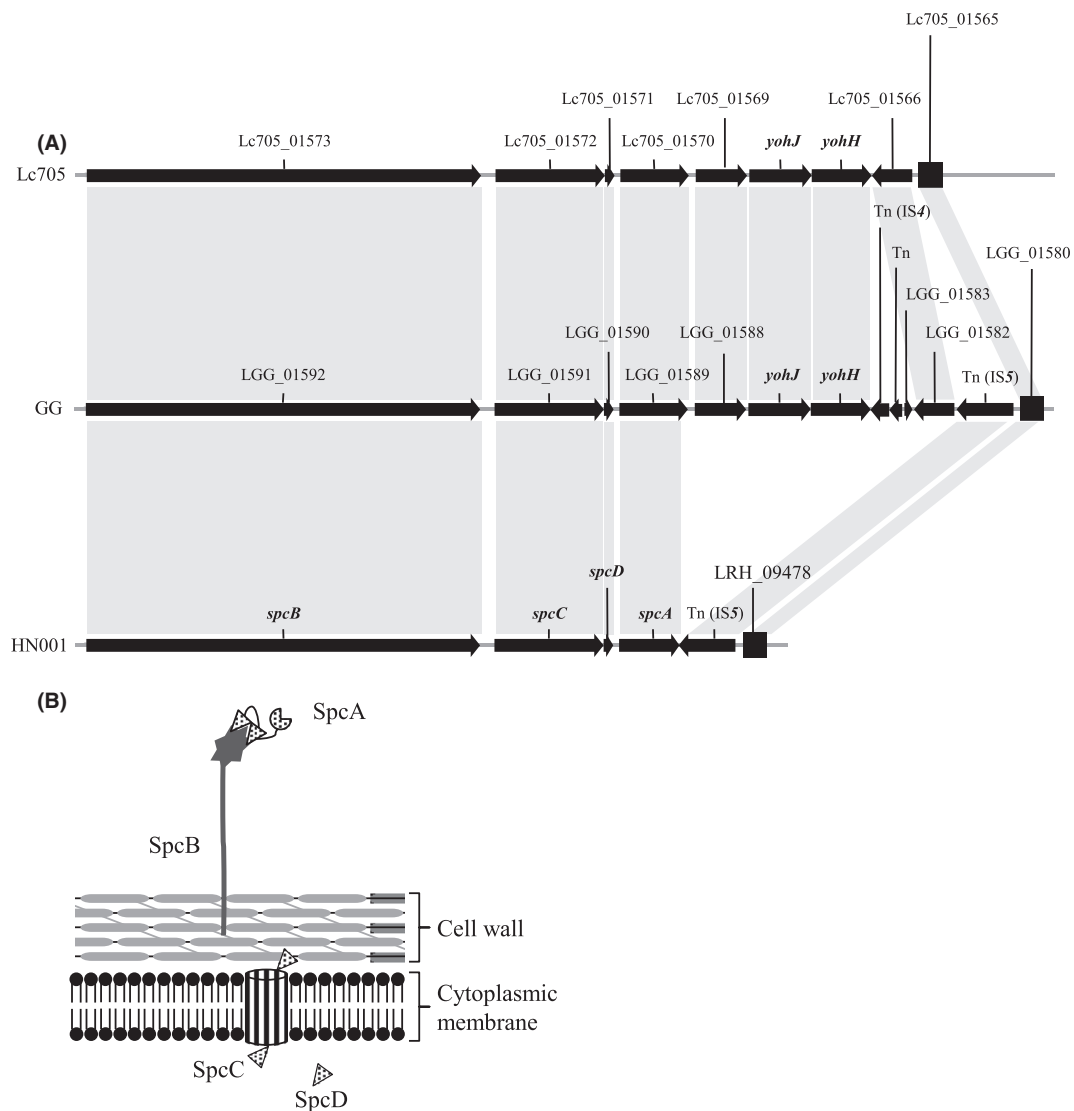
We report a discovery of two extracellular proteins that appear to form a complex on the surface of the probiotic bacterium *L. rhamnosus* HN001. This complex was identified by affinity screenings of an *L. rhamnosus* HN001 phage display library. In the first screen, the cell-binding protein SpcA was identified using the *L. rhamnosus* HN001 cells as bait. Purified SpcA was, in turn, used to identify its ligand (or docking protein), SpcB, a large cell-wall-anchored adhesin/fibril-like protein. Both proteins are unique to *L. rhamnosus*, with only a distant homolog of docking protein found in the genome of *L. parafarraginis*. These findings are consistent with the species specificity in cell-binding assays using SpcA-displaying PPs.

Mapping of binding domains determined that the two bacterial immunoglobulin-like domains 3 (Big-3) of SpcA interacted with an N-terminal high-complexity domain of SpcB. SpcA also contains a domain that showed the best

match in a 3D alignment with the bacterial PR-1-like domain, but also aligns weakly with  $\beta$ -glycanases.

The SpcA-docking protein SpcB has a signal sequence motif and organization typical of Gram-positive glycosylated adhesins, exemplified by the GspB adhesin of *Streptococcus gordonii* that binds to human platelets, and a mucus-binding fibril-like adhesin of *Streptococcus parasanguinis*, Fap1 (Wu et al. 1998; Bensing et al. 2007). Even though the signal sequence of SpcB has characteristics typical of the glycosylated Ser-rich adhesins, the gene organization surrounding the *spcB* ORF is very different from that of *gspB* and *fap1*. The *gspB* and *fap1* genes are in the same cluster/operon with two glycosylase-encoding genes and genes encoding for a secondary/alternative Sec

translocon complex (SecA2, SecY2 and in some cases SecE2 and SecG2) (Rigel and Braunstein 2008; Zhou and Wu, 2009). The secondary translocon, either in combination with or independently of the standard translocon (SecA, SecYEG), is required for export of GspB and Fap1 from their respective organisms. The same gene cluster also typically encodes two or more transmembrane or cytosolic accessory proteins that either facilitate or are required for the secretion of glycosylated precursors (Rigel and Braunstein 2008; Zhou and Wu 2009). In contrast, the *spcB*–*spcA* cluster in HN001 includes only two other genes: *spcC*, encoding a transmembrane protein containing two Big-3 domains; and *spcD*, encoding a small cytosolic protein containing a single Big-3 domain (Fig. 6A).



**Figure 6.** (A) Organization of *spcB*–*spcA* gene cluster in *Lactobacillus rhamnosus* strains Lc 705, GG, and HN001 and GG. (B) Model of the SpcB–SpcA complex and position of SpcC and SpcD.



Analysis of sequence downstream sequence of ORF *spcA* in HN001 revealed the presence of an ORF, transcribed in the opposite direction, encoding a putative transposase within a predicted transposable element, possibly belonging to insertional element of IS5 family (IS5). Using the Find-Term bacterial terminators prediction algorithm (<http://www.softberry.ru/berry.phtml>), two terminator sequences were detected: one upstream of putative transcriptional start of *spcB* and one bordering the putative IS5 element. No terminator sequences were detected using the same prediction software between the *spcB*, *spcC*, *spcD*, and *spcA* ORFs (Jankovic 2009).

This cluster of ORFs (*spcB*–*spcA*) is conserved in the GG genome (LGG\_01592–LGG\_01589) and the genomes of other *L. rhamnosus* strains (Kankainen et al. 2009; Morita et al. 2009; Yu et al. 2011; Pittet et al. 2012; Prajapati et al. 2012). However, starting from the stop codon after the residue 503 of *spcA*, the HN001 sequence contains an IS5-like element, whereas GG (ATCC 53103), Lc 705, and ATCC 8530 contain the complete *spcA* ORF, followed by two additional ORFs that encode two predicted glycosyltransferases (GtfA family) that are normally found in the same gene cluster as large KxYKxGKxW-signal-sequence-containing adhesins and are required for glycosylation of Fap1 adhesin (Zhou and Wu 2009). Downstream of *spcA* are three IS elements in GG, but none in Lc 705. Interestingly, the *spcA*-inserted IS5-like element of HN001 is at the end of the NZ\_ABWJ01000002.1 contig; to determine the sequence beyond the end of this contig, we used a series of primers corresponding to *spcA*-downstream sequence in GG. This analysis detected a deletion that removed seven ORFs between *spcA* and the IS5 element in HN001 that corresponds to the second IS element in GG, downstream from the *spcA* (Fig. 6).

Even though HN001 lacks the two GtfA family glycosyltransferases, it contains four additional GtfA homologs that are highly conserved to both missing glycosyltransferases. Therefore, it is possible that SpcB is still glycosylated in HN001, despite the deletion. Whatever possible differences in modification between HN001 and GG, they do not affect binding of SpcA, as shown in Figure 3.

The alternative Sec system has not been annotated in any of the sequenced *L. rhamnosus* genomes. Given that *L. rhamnosus* has only a single complement of genes annotated as *secA* and *secYEG*, it is expected that SpcB is exported via the standard SecA–SecYEG system. Proteins encoded by the two ORFs between SpcB and SpcA could be accessory proteins involved in secretion of glycosylated SpcB.

The reported proteomic analysis of *L. rhamnosus* GG surface proteins identified, using mass spectroscopy, 105 different peptides from SpcB (LGG\_01592) (Savijoki et al. 2011). This protein is therefore produced in *L. rhamnosus* GG and targeted to the cell wall. The mass fingerprinting

approach used in Savijoki et al. (2011) cannot assign peptides to a protein if they are glycosylated; therefore, many proteolytic peptides derived SpcB in *L. rhamnosus* GG must be unmodified (Lebeer et al. 2012).

The Ala-Ser-rich region of SpcB has some characteristics of  $\alpha$ -helical membrane-inserting bacterial proteins. The PHYRE server predicted helical secondary structure along this region, but at the same time predicted, with high probability, disordered sequence. Given that we demonstrated that the Ala-Ser-rich region of SpcB (residues 401–1209) can insert into phospholipid bilayers (Jankovic et al. 2007), this portion of the molecule could target *L. rhamnosus* to phospholipid bilayers, including membranes of host epithelial cells. Alternatively, depending on the extent of glycosylation and the number of copies on the surface of *L. rhamnosus* cells, SpcB could regulate hydrophobicity or hydration of the cell surface.

Using a phage display approach, we identified a protein complex, representing a putative fibril-like-docking protein SpcB and attached soluble protein, SpcA, displaying a PR-1-like domain (Fig. 6B). SpcB–SpcA complex, due to the PR-1-like domain of SpcA and membrane-penetrating hydrophobic region of SpcB, has potential to mediate specific protein–protein interactions, membrane-insertion, or cell-surface-hydrophobicity regulation functions, all of which are pertinent to interaction with the environment and host.

## Acknowledgments

This work was funded by the Institute of Molecular Bio-Sciences (Massey University), Palmerston North Medical Research Foundation, Anonymous Donor, New Zealand Foundation for Research and Technology (contract number C03X0701), and Fonterra Co-operative Group Ltd. D. G. was a recipient of a Bright Future Enterprise Ph.D. Scholarship (Ministry of Science and Innovation, New Zealand). Yun Wu is acknowledged for technical help in construction of the vector pYW01 and James Dekker for critical reading of the manuscript.

*Author contributions:* D. G. has constructed the vector and HN001 phage display library, identified SpcA by library screening, and did the domain SpcA domain mapping. W. W. has constructed a  $\Delta$ spcA mutant of HN001, did SpcA PP and SpcB PP-binding assays, and screened the HN001 phage display library for proteins that bind to SpcA. M. A. C. contributed advice,  $\Delta$ weeE and  $\Delta$ dltD HN001 mutants and tools for construction of the  $\Delta$ spcA mutant. J. R. directed the research and has written the manuscript (together with W. W. and D. G., with contribution from M. A. C.).

## Conflict of Interest

None declared.

## References

- Altermann, E., W. M. Russell, M. A. Azcarate-Peril, R. Barrangou, B. L. Buck, O. McAuliffe, et al. 2005. Complete genome sequence of the probiotic lactic acid bacterium *Lactobacillus acidophilus* NCFM. Proc. Natl. Acad. Sci. USA 102:3906–3912.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.
- Armitage, R. J., W. C. Fanslow, L. Strockbine, T. A. Sato, K. N. Clifford, B. M. Macduff, et al. 1992. Molecular and biological characterization of a murine ligand for CD40. Nature 357:80–82.
- Ausmees, N., K. Jacobsson, and M. Lindberg. 2001. A unipolarly located, cell-surface-associated agglutinin, RapA, belongs to a family of Rhizobium-adhering proteins (Rap) in *Rhizobium leguminosarum* bv. trifolii. Microbiology 147:549–559.
- Baltrus, D. A., M. T. Nishimura, A. Romanchuk, J. H. Chang, M. S. Mukhtar, K. Cherkis, et al. 2011. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. PLoS Pathog. 7:e1002132.
- Beckmann, C., J. D. Waggoner, T. O. Harris, G. S. Tamura, and C. E. Rubens. 2002. Identification of novel adhesins from Group B streptococci by use of phage display reveals that C5a peptidase mediates fibronectin binding. Infect. Immun. 70:2869–2876.
- Beekwilder, J., J. Rakonjac, M. Jongsma, and D. Bosch. 1999. A phagemid vector using the *E. coli* phage shock promoter facilitates phage display of toxic proteins. Gene 228:23–31.
- Bensing, B. A., I. R. Siboo, and P. M. Sullam. 2007. Glycine residues in the hydrophobic core of the GspB signal sequence route export toward the accessory Sec pathway. J. Bacteriol. 189:3846–3854.
- Boltana, S., F. Reyes-Lopez, D. Morera, F. Goetz, and S. A. MacKenzie. 2011. Divergent responses to peptidoglycans derived from different *E. coli* serotypes influence inflammatory outcome in trout, *Oncorhynchus mykiss*, macrophages. BMC Genomics 12:34.
- Bron, P. A., P. van Baarlen, and M. Kleerebezem. 2012. Emerging molecular insights into the interaction between probiotics and the host intestinal mucosa. Nat. Rev. Microbiol. 10:66–78.
- Buck, B. L., E. Altermann, T. Svingerud, and T. R. Klaenhammer. 2005. Functional analysis of putative adhesion factors in *Lactobacillus acidophilus* NCFM. Appl. Environ. Microbiol. 71:8344–8351.
- Buts, L., J. Bouckaert, E. De Genst, R. Loris, S. Oscarson, M. Lahmann, et al. 2003. The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds N-acetylglucosamine. Mol. Microbiol. 49:705–715.
- Carroll, R. K., S. B. Beres, I. Sitkiewicz, L. Peterson, R. K. Matsunami, D. A. Engler, et al. 2011. Evolution of diversity in epidemics revealed by analysis of the human bacterial pathogen group A *Streptococcus*. Epidemics 3:159–170.
- Castaldo, C., V. Vastano, R. A. Siciliano, M. Candela, M. Vici, L. Muscariello, et al. 2009. Surface displaced alpha-enolase of *Lactobacillus plantarum* is a fibronectin binding protein. Microb. Cell Fact. 8:14.
- Collado, M. C., L. Jalonen, J. Meriluoto, and S. Salminen. 2006. Protection mechanism of probiotic combination against human pathogens: in vitro adhesion to human intestinal mucus. Asia Pac. J. Clin. Nutr. 15:570–575.
- Collado, M. C., J. Meriluoto, and S. Salminen. 2007. Development of new probiotics by strain combinations: is it possible to improve the adhesion to intestinal mucus? J. Dairy Sci. 90:2710–2716.
- Cross, M. L., R. R. Mortensen, J. Kudsk, and H. S. Gill. 2002. Dietary intake of *Lactobacillus rhamnosus* HNOO1 enhances production of both Th1 and Th2 cytokines in antigen-primed mice. Med. Microbiol. Immunol. (Berl.) 191:49–53.
- Dave, M., P. D. Higgins, S. Middha, and K. P. Rioux. 2012. The human gut microbiome: current knowledge, challenges, and future directions. Transl. Res. 160:246–257.
- Dekker, J., M. Collett, J. Prasad, and P. Gopal. 2007. Functionality of probiotics – potential for product development. Forum Nutr. 60:196–208.
- Elberse, K., S. Witteveen, H. van der Heide, I. van de Pol, C. Schot, A. van der Ende, et al. 2011. Sequence diversity within the capsular genes of *Streptococcus pneumoniae* serogroup 6 and 19. PLoS ONE 6:e25018.
- Eskan, M. A., G. Hajishengallis, and D. F. Kinane. 2007. Differential activation of human gingival epithelial cells and monocytes by *Porphyromonas gingivalis* fimbriae. Infect. Immun. 75:892–898.
- Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, et al. 2006. Pfam: clans, web tools and services. Nucleic Acids Res. 34:D247–D251.
- Forrer, P., S. Jung, and A. Pluckthun. 1999. Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins. Curr. Opin. Struct. Biol. 9:514–520.
- Fraser, J. S., K. L. Maxwell, and A. R. Davidson. 2007. Immunoglobulin-like domains on bacteriophage: weapons of modest damage? Curr. Opin. Microbiol. 10:382–387.
- Geer, L. Y., M. Domrachev, D. J. Lipman, and S. H. Bryant. 2002. CDART: protein homology by domain architecture. Genome Res. 12:1619–1623.
- Gill, H. S., Q. Shu, H. Lin, K. J. Rutherford, and M. L. Cross. 2001. Protection against translocating *Salmonella*

- typhimurium* infection in mice by feeding the immunoenhancing probiotic *Lactobacillus rhamnosus* strain HN001. *Med. Microbiol. Immunol. (Berl.)* 190:97–104.
- Gillen, C. M., H. S. Courtney, K. Schulze, M. Rohde, M. R. Wilson, A. M. Timmer, et al. 2008. Opacity factor activity and epithelial cell binding by the serum opacity factor protein of *Streptococcus pyogenes* are functionally discrete. *J. Biol. Chem.* 283:6359–6366.
- Grandel, U., and F. Grimminger. 2003. Endothelial responses to bacterial toxins in sepsis. *Crit. Rev. Immunol.* 23:267–299.
- Grice, E. A., and J. A. Segre. 2011. The skin microbiome. *Nat. Rev. Microbiol.* 9:244–253.
- Hao, W., J. H. Ma, K. Warren, R. S. Tsang, D. E. Low, F. B. Jamieson, et al. 2011. Extensive genomic variation within clonal complexes of *Neisseria meningitidis*. *Genome Biol. Evol.* 3:1406–1418.
- Hultgren, S. J., F. Jacob-Dubuisson, C. H. Jones, and C. I. Branden. 1993. PapD and superfamily of periplasmic immunoglobulin-like pilus chaperones. *Adv. Protein Chem.* 44:99–123.
- Hummel, S., K. Veltman, C. Cichon, U. Sonnenborn, and M. A. Schmidt. 2012. Differential targeting of the E-cadherin/b-catenin complex by Gram-positive probiotic *Lactobacilli* improves epithelial barrier function. *Appl. Environ. Microbiol.* 78:1140–1147.
- Jacobsson, K., A. Rosander, J. Bjerketorp, and L. Frykberg. 2003. Shotgun phage display – selection for bacterial receptors or other exported proteins. *Biol. Proced. Online* 5:123–135.
- Jankovic, D. 2009. Direct selection and phage display of *Lactobacillus rhamnosus* HN001 secretome. Massey University, Palmerston North.
- Jankovic, D., M. A. Collett, M. W. Lubbers, and J. Rakonjac. 2007. Direct selection and phage display of a Gram-positive secretome. *Genome Biol.* 8:R266.
- Jenkinson, H. F. 2011. Beyond the oral microbiome. *Environ. Microbiol.* 13:3077–3087.
- Kalliomaki, M., S. Salminen, T. Poussa, H. Arvilommi, and E. Isolauri. 2003. Probiotics and prevention of atopic disease: 4-year follow-up of a randomised placebo-controlled trial. *Lancet* 361:1869–1871.
- Kankainen, M., L. Paulin, S. Tynkkyinen, I. von Ossowski, J. Reunanen, P. Partanen, et al. 2009. Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. *Proc. Natl. Acad. Sci. USA* 106:17193–17198.
- Kelley, L. A., and M. J. Sternberg. 2009. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* 4:363–371.
- Kinross, J. M., A. W. Darzi, and J. K. Nicholson. 2011. Gut microbiome-host interactions in health and disease. *Genome Med.* 3:14.
- Kleerebezem, M., P. Hols, E. Bernard, T. Rolain, M. Zhou, R. J. Siezen, et al. 2010. The extracellular biology of the lactobacilli. *FEMS Microbiol. Rev.* 34:199–230.
- Kolenbrander, P. E., R. J. Palmer, Jr., A. H. Rickard, N. S. Jakubovics, N. I. Chalmers, and P. I. Diaz. 2006. Bacterial interactions and successions during plaque development. *Periodontol.* 2000:47–79.
- Konstantinov, S. R., H. Smidt, W. M. de Vos, S. C. Bruijns, S. K. Singh, F. Valence, et al. 2008. S layer protein A of *Lactobacillus acidophilus* NCFM regulates immature dendritic cell and T cell functions. *Proc. Natl. Acad. Sci. USA* 105:19474–19479.
- Kowalchuk, G. A., A. G. Speksnijder, K. Zhang, R. M. Goodman, and J. A. van Veen. 2007. Finding the needles in the metagenome haystack. *Microb. Ecol.* 53:475–485.
- Kulkarni, R., B. K. Dhakal, E. S. Slechta, Z. Kurtz, M. A. Mulvey, and D. G. Thanassi. 2009. Roles of putative type II secretion and type IV pilus systems in the virulence of uropathogenic *Escherichia coli*. *PLoS ONE* 4:e4752.
- Laemmli, U. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685.
- Lahtinen, S. J., S. Forssten, J. Aakko, L. Granlund, N. Rautonen, S. Salminen, et al. 2012. Probiotic cheese containing *Lactobacillus rhamnosus* HN001 and *Lactobacillus acidophilus* NCFM(R) modifies subpopulations of fecal lactobacilli and *Clostridium difficile* in the elderly. *Age (Dordr)* 34:133–143.
- Lebeer, S., I. J. Claes, C. I. Balog, G. Schoofs, T. L. Verhoeven, K. Nys, et al. 2012. The major secreted protein Msp1/p75 is O-glycosylated in *Lactobacillus rhamnosus* GG. *Microb. Cell Fact.* 11:15.
- Lilic, M., M. Vujanac, and C. E. Stebbins. 2006. A common structural motif in the binding of virulence factors to bacterial secretion chaperones. *Mol. Cell* 21:653–664.
- Lin, P. W., L. E. Myers, L. Ray, S. C. Song, T. R. Nasr, A. J. Berardinelli, et al. 2009. *Lactobacillus rhamnosus* blocks inflammatory signaling in vivo via reactive oxygen species generation. *Free Radic. Biol. Med.* 47:1205–1211.
- Luck, S. N., L. Badea, V. Bennett-Wood, R. Robins-Browne, and E. L. Hartland. 2006. Contribution of FliC to epithelial cell invasion by enterohemorrhagic *Escherichia coli* O113:H21. *Infect. Immun.* 74:6999–7004.
- Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, et al. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39:D225–D229.
- McGhie, E. J., L. C. Brawn, P. J. Hume, D. Humphreys, and V. Koronakis. 2009. Salmonella takes control: effector-driven manipulation of the host. *Curr. Opin. Microbiol.* 12:117–124.
- Mizuki, N., and M. Kasahara. 1992. Mouse submandibular glands express an androgen-regulated transcript encoding an acidic epididymal glycoprotein-like molecule. *Mol. Cell. Endocrinol.* 89:25–32.
- Morita, H., H. Toh, K. Oshima, M. Murakami, T. D. Taylor, S. Igimi, et al. 2009. Complete genome sequence of the

- probiotic *Lactobacillus rhamnosus* ATCC 53103. *J. Bacteriol.* 191:7630–7631.
- Mullen, L. M., S. P. Nair, J. M. Ward, A. N. Rycroft, and B. Henderson. 2006. Phage display in the study of infectious diseases. *Trends Microbiol.* 14:141–147.
- Munoz-Provencio, D., and V. Monedero. 2011. Shotgun phage display of *Lactobacillus casei* BL23 against collagen and fibronectin. *J. Microbiol. Biotechnol.* 21:197–203.
- Munoz-Provencio, D., G. Perez-Martinez, and V. Monedero. 2010. Characterization of a fibronectin-binding protein from *Lactobacillus casei* BL23. *J. Appl. Microbiol.* 108:1050–1059.
- Patti, J. M., B. L. Allen, M. J. McGavin, and M. Hook. 1994. MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu. Rev. Microbiol.* 48:585–617.
- Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8:785–786.
- van Pijkeren, J. P., C. Canchaya, K. A. Ryan, Y. Li, M. J. Claesson, B. Sheil, et al. 2006. Comparative and functional analysis of sortase-dependent proteins in the predicted secretome of *Lactobacillus salivarius* UCC118. *Appl. Environ. Microbiol.* 72:4143–4153.
- Pittet, V., E. Ewen, B. R. Bushell, and B. Ziola. 2012. Genome sequence of *Lactobacillus rhamnosus* ATCC 8530. *J. Bacteriol.* 194:726.
- Prajapati, J. B., C. D. Khedkar, J. Chitra, S. Suja, V. Mishra, V. Sreeja, et al. 2012. Whole-genome shotgun sequencing of *Lactobacillus rhamnosus* MTCC 5462, a strain with probiotic potential. *J. Bacteriol.* 194:1264–1265.
- Preidis, G. A., C. Hill, R. L. Guerrant, B. S. Ramakrishna, G. W. Tannock, and J. Versalovic. 2011. Probiotics, enteric and diarrheal diseases, and global health. *Gastroenterology* 140:8–14.
- Rakonjac, J., N. J. Bennett, J. Spagnuolo, D. Gagic, and M. Russel. 2011. Filamentous bacteriophage: biology, phage display and nanotechnology applications. *Curr. Issues Mol. Biol.* 13:51–75.
- Rigel, N. W., and M. Braunstein. 2008. A new twist on an old pathway – accessory Sec [corrected] systems. *Mol. Microbiol.* 69:291–302.
- Rimoldi, M., M. Chieppa, V. Salucci, F. Avogadri, A. Sonzogni, G. M. Sampietro, et al. 2005. Intestinal immune homeostasis is regulated by the crosstalk between epithelial cells and dendritic cells. *Nat. Immunol.* 6:507–514.
- Russel, M. 1993. Protein-protein interactions during filamentous phage assembly. *J. Mol. Biol.* 231:689–697.
- Sambrook, J., and D. W. Russell. 2001. *Molecular cloning: a laboratory manual*. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Sansonetti, P. J. 2004. War and peace at mucosal surfaces. *Nat. Rev. Immunol.* 4:953–964.
- Sansonetti, P. J. 2011. To be or not to be a pathogen: that is the mucosally relevant question. *Mucosal Immunol.* 4:8–14.
- Sansonetti, P. J., and R. Medzhitov. 2009. Learning tolerance while fighting ignorance. *Cell* 138:416–420.
- Savijoki, K., N. Lietzen, M. Kankainen, T. Alatossava, K. Koskeniemi, P. Varmanen, et al. 2011. Comparative proteome cataloging of *Lactobacillus rhamnosus* strains GG and Lc705. *J. Proteome Res.* 10:3460–3473.
- Schneewind, O., P. Model, and V. A. Fischetti. 1992. Sorting of protein A to the staphylococcal cell wall. *Cell* 70:267–281.
- Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc. Natl. Acad. Sci. USA* 95:5857–5864.
- Selbach, M., and S. Backert. 2005. Cortactin: an Achilles' heel of the actin cytoskeleton targeted by pathogens. *Trends Microbiol.* 13:181–189.
- Sheih, Y. H., B. L. Chiang, L. H. Wang, C. K. Liao, and H. S. Gill. 2001. Systemic immunity-enhancing effects in healthy subjects following dietary consumption of the lactic acid bacterium *Lactobacillus rhamnosus* HN001. *J. Am. Coll. Nutr.* 20:149–156.
- Shu, Q., and H. S. Gill. 2002. Immune protection mediated by the probiotic *Lactobacillus rhamnosus* HN001 (DR20) against *Escherichia coli* O157:H7 infection in mice. *FEMS Immunol. Med. Microbiol.* 34:59–64.
- Sidhu, S. S., W. J. Fairbrother, and K. Deshayes. 2003. Exploring protein-protein interactions with phage display. *ChemBioChem* 4:14–25.
- Song, C., A. Kumar, and M. Saleh. 2009. Bioinformatic comparison of bacterial secretomes. *Genomics Proteomics Bioinformatics* 7:37–46.
- Swanson, P. A., II, A. Kumar, S. Samarin, M. Vijay-Kumar, K. Kundu, N. Murthy, et al. 2011. Enteric commensal bacteria potentiate epithelial restitution via reactive oxygen species-mediated inactivation of focal adhesion kinase phosphatases. *Proc. Natl. Acad. Sci. USA* 108:8803–8808.
- Takai, T., and M. Ono. 2001. Activating and inhibitory nature of the murine paired immunoglobulin-like receptor family. *Immunol. Rev.* 181:215–222.
- Thomas, D. J., R. J. Husmann, M. Villamar, T. R. Winship, R. H. Buck, and F. A. Zuckermann. 2011. *Lactobacillus rhamnosus* HN001 attenuates allergy development in a pig model. *PLoS ONE* 6:e16577.
- Timmer, A. M., S. A. Kristian, V. Datta, A. Jeng, C. M. Gillen, M. J. Walker, et al. 2006. Serum opacity factor promotes group A streptococcal epithelial cell invasion and virulence. *Mol. Microbiol.* 62:15–25.
- Toba, T., R. Virkola, B. Westerlund, Y. Bjorkman, J. Sillanpaa, T. Vartio, et al. 1995. A collagen-binding S-layer protein in



- Lactobacillus crispatus*. Appl. Environ. Microbiol. 61: 2467–2471.
- Tuanyok, A., J. K. Stone, M. Mayo, M. Kaestli, J. Gruendike, S. Georgia, et al. 2012. The genetic and molecular basis of O-antigenic diversity in *Burkholderia pseudomallei* lipopolysaccharide. PLoS Negl. Trop. Dis. 6: e1453.
- Walter, J., and R. Ley. 2011. The human gut microbiome: ecology and recent evolutionary changes. Annu. Rev. Microbiol. 65:411–429.
- Wells, J. M. 2011. Immunomodulatory mechanisms of lactobacilli. Microb. Cell Fact. 10(Suppl. 1):S17.
- Wickens, K., P. Black, T. V. Stanley, E. Mitchell, C. Barthow, P. Fitzharris, et al. 2012. A protective effect of *Lactobacillus rhamnosus* HN001 against eczema in the first 2 years of life persists to age 4 years. Clin. Exp. Allergy 42: 1071–1079.
- Wu, H., K. P. Mintz, M. Ladha, and P. M. Fives-Taylor. 1998. Isolation and characterization of Fap1, a fimbriae-associated adhesin of *Streptococcus parasanguis* FW213. Mol. Microbiol. 28:487–500.
- Yu, B., F. Su, L. Wang, B. Zhao, J. Qin, C. Ma, et al. 2011. Genome sequence of *Lactobacillus rhamnosus* strain CASL, an efficient L-lactic acid producer from cheap substrate cassava. J. Bacteriol. 193:7013–7014.
- Zhou, M., D. Theunissen, M. Wels, and R. J. Siezen. 2010. LAB-Secretome: a genome-scale comparative analysis of the predicted extracellular and surface-associated proteins of Lactic Acid Bacteria. BMC Genomics 11:651.
- Zhou, M., and H. Wu. 2009. Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. Microbiology 155:317–327.
- Zwick, M. G., J. Shen, and J. K. Scott. 1998. Phage-displayed peptide libraries. Curr. Opin. Biotechnol. 9:427–436.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** The plasmid profiles over three rounds of affinity-screening (panning) of the HN001 phage display library using HN001 as bait. Lanes: pYW01, replicative form (covalently closed circular double-stranded DNA; cccDNA) of the phagemid vector pYW01 used to construct the phage display library; Round I; Round II; Round III; purified pools of the recombinant library (cccDNA) after indicated rounds of panning. Three black arrows indicate the DNA bands of enriched phagemids.

**Figure S2.** Construction of the HN001  $\Delta$ *spcA* mutant. (A) Vector used for construction of double cross-over mutant. (B) Map of the *spcA* deletion mutant and the wild-type *spcA* locus in HN001. (C) Confirmation Southern blotting using upstream and downstream probes (indicated in B).

**Figure S3.** The plasmid profiles over four rounds of affinity-screening (panning) of the HN001 phage display library using purified MBP-SpcA fusion protein as bait. The panning was carried out over four rounds of affinity selection and amplification (I–IV) to enrich the specific binders. From the second to the fourth rounds of panning, MBP and BSA were also included as control baits. pYW01 represents the phagemid used to construct the phage display library.

**Table S1.** Bacterial strains.

**Table S2.** Plasmids and phage.

**Table S3.** Oligonucleotides.

**Data S1.** Experimental procedures.